DOI：10.13998/j.cnki.issn1002-1248.23-0428

全文本引文分析研究综述

李　点，文庭孝*，许林勇*

（中南大学 生命科学学院，长沙　410006）

摘　要：［目的 / 意义］梳理全文本引文分析当前发展现状，明确国内外全文本引文分析的不同侧重，剖析全文本引文分析面临的挑战以及未来可能的发展方向。［方法 / 过程］通过回顾国内外关于全文本引文分析相关研究性论文结果，将全文本引文分析研究内容总结为引用位置分析、引用强度分析、引用语境分析，并对全文本引文分析在跨学科、科学评价等领域上的应用进行分析，并总结国内外全文本引文分析差异。［结果 / 结论］通过文献梳理，发现全文本引用分析目前在引用位置上的研究与章节结构有重要联系。引用强度分析与引用位置分析结合紧密，并注重开发新的计算方法。引用语境上，全文本引用分析的热点主要是情感分析。应用集中在交叉学科测度以及科学评价体系构建上。本研究还针对全文本引文分析面临的挑战以及今后的发展方向做出了预测。

关键词：全文本分析；引文分析；情感分析；引用位置分析

中图分类号：G256　　　**文献标识码：**A　　　**文章编号：**1002-1248（2023）09-0015-13

引用本文：李点，文庭孝，许林勇. 全文本引文分析研究综述［J］. 农业图书情报学报，2023，35（9）：15-27.

1 引　言

引文分析是文献计量学中的重要研究领域，通过分析文献中的引用关系，研究文献的引用行为和学术影响力。20 世纪 90 年代以来，研究者们开始就全引文对文献进行研究分析，传统的引用分析主要基于参考文献列表和作者引用标识，但这种方法存在一些问题，如无法识别全文本引用和保证引用的准确性，也较少涉及引用内容的分析。随着文本挖掘以及机器学习技术的进一步发展，全文开放数据库陆续成立并对公众开放，全文本引文分析方法得到了快速发展，成为文献计量学研究的重要方向。

全文本引文分析是一种通过分析文献中的文本信息，自动识别文献间的引用关系的分析方法。相较于传统的引用分析，全文本引文分析更加准确和全面，能够帮助研究者更好地了解文献的引用关系和学术影响。全文本引文分析主要涉及机器学习和自然语言处理技术，包括文本特征提取、聚类分析、关键词抽取、主题模型和情感分析等。

目前，全文本引文分析已经在文献计量学、科学评价、学科发展趋势分析等领域得到了广泛的应用。

收稿日期：2023-06-30

基金项目：国家社科基金项目“基于人体信息计量的居民健康指数构建与精准画像研究”（2BTQ052）

作者简介：李点（1999- ），女，硕士，研究方向为文献计量学

*通信作者：文庭孝（1975- ），男，博士，教授，博士生导师，研究方向为信息计量与科学评价。Email：wtxsomebody@aliyun.com。许林勇（1977- ），男，博士，硕士生导师，副教授，研究方向为生物信息与医学统计。Email：xybms@163.com

DOI：10.13998/j.cnki.issn1002-1248.23-0428

然而，随着数据量的增加，全文本引文分析也面临着一些挑战，如数据质量、精度和效率、跨语言引用分析等问题。因此，需要对全文本引文分析进行梳理，以更好地推动全文本引文分析发展。

在正式检索前对全文本研究领域文章进行初步检索，检索词初步设定为全文本引文分析及其同义词，并在检索过程中根据检索所得文章的具体研究对象的细分对关键词进行补充完善。最终本文以关键词（全文本分析、引用强度分析、引用情感分析、引用位置分析、引用动机分析、引用分析评价体系、跨学科引文分析、引文分析应用等）的英文单词构造相关检索式，关键词关系为 OR，在 Web of Science 中进行主题检索，设定文献类型 Article，筛选早期高被引文章，并明确时间跨度为 2000—2023 年，共得到 135 篇非综述类型相关文章，根据文章题目以及摘要进行人工筛选得到相关文献共 72 篇。与此同时用相应中文名词构造关键词检索式并在中国知网（CNKI）上进行主题检索，初步得到文献共 128 篇，同样对文章标题以及摘要进行人工筛选，最终得到中文文献 53 篇。最后对中英文文章进行综合，去除同作者相似文章。如此，共检得文献 125 篇以进行下文分析。

在检索所得文献的基础上，对全文本引文分析进行系统梳理。根据检索所得文章的具体研究内容归类主要分为对于全文本内容具体分析以及全文本分析的具体应用。故可将文章结构主要分为以下 3 部分。第一部分全文本引文分析的研究内容，又可分为引用位置分析，引用强度分析，引用语境分析（主要依据各篇文章所涉及关键词以及词语重复程度确定分为以上三细点进行分析）。引用位置、强度、语境可以看成是引用具体内容的 3 个不同方面，引用所处的文章结构不同、引用次数不同、引用所带情感选用词语等语境不同，引用的效果均会有所不同，具体展开见下文。第二部分全文本引文分析的主要应用，由跨学科研究，科学评价以及其他领域应用 3 个部分组成，第三部分对全文本引文分析研究进行展望，分析面临的挑战以及未来可能的发展趋势。

2 全文本引文分析的兴起与发展

在相当长一段时间里，由于论文数据规模不大，全文文本以及数据难以获取等原因，全文本相关分析研究以及相关引文研究处于早期阶段，研究数量稀少。进入本世纪，特别是电脑计算机技术的大规模应用以来，开放获取运动逐渐兴起，自然语言处理以及文本挖掘等技术快速兴起与发展，更方便获取学术论文全文数据进行分析，引文分析在内的全文本分析研究重新进入众多研究者的视野。BIRNDL、BIR 等相关国际专题会议近年来兴起，相关研究者致力于信息检索与文献计量以及自然语言处理等领域交叉研究，共同探索全文本内容分析本身以及其在文献计量、信息检索等领域的应用。目前全文本引文分析研究内容主要集中在引用位置分析、引用强度分析、引用语境分析以及全文本引文分析应用等方面[1]。

3 全文本引文分析的研究内容

3.1 引用位置分析

全文本数据的逐渐开放，研究者们可以不局限于文章末尾的参考文献数据，同时可以进一步定位参考文献在文章中被引用的具体位置。关于引文位置方向的研究，可以总结为定性与定量两个方面。

（1）引用位置定性分析。早在 2000 年，何荣利等[2]就对 300 篇文章及其 7 857 篇引文在引言、本来、结论 3 个部分的分布数量进行了研究，得出引文作用与所在文章部分有关；2010 年李燕[3]认为简单根据引文和被引来对科技期刊进行排名不太准确，并以此为契机进行研究，最终发现引文在文中位置分布能影响其引文效用；刘茜等[4]，出于对 SCI 文章评价体现中“引文均等”这一假设的疑惑，对文章引文动机认知理论进行梳理并分类，最终研究结果显示被引文献在施引文章的引用位置是随时间的变化进行有规律改变。

（2）引用位置定量分析。2013 年 HALEVI 等[5]对

Informetrics 杂志 2007 年发表的文章进行分析，分析了引文在文章不同部分中的使用情况，并通过矩阵来进行定量部分的计算；2016 年胡志刚[6]利用可视化对引用位置的分布进行直观展示，对不同年份的引用位置和不同被引次数引文的位置进行比较，并对二者的特点归纳总结；2017 年章成志等[7]以学术专著为研究对象，对其进行人工标注并从引文内容位置分布等角度分析学术专著的引文行为。这一研究成果将全文本引用分析的研究对象从常用的研究性论文扩大至学术专著，为以后对其他学术成果格式例如专利的研究起到了促进作用。张琳等[8]基于论文的不同章节的引文标注位置对学科交叉度进行分析，结果表明引言、讨论、方法和结果 4 个章节的学科交叉程度依次降低[9]，并在此基础上提出了加权 TD 学科交叉测度指标，为引用位置分析进一步应用提供了动力。

3.2 引用强度分析

引用强度分析主要是指引文在文献中的被引用的次数及其再不同段落的分布情况。一般来说，被引次数的多少与引用强度的大小成正比关系。但这其中也包含了引文在文献中被多次引用时在不同位置的分布差异状况，分布不同，引用强度的大小也会不同。

（1）单强度分析。2013 年 DING 等[10]提出了两种不同的引用强度计算方法，分别命名为 Count One 和 Count X，并发现使用的计数不同导致的分析结果也有所差异，在实际的研究中需要根据自己的研究目的进行方法辨析。胡志刚等[11]研究分析了引文文章数量和引用频次这两者之间的联系，并在此基础上提出了一种新的加权计算引文总被引次数的方法，该方法在科学评价方面具有重要价值。胡志刚长期关注该领域的发展[6]，其 2017 年发表的文章[12]中，对“怎样的引文更容易多次引用，怎样的施引文献更容易进行多次引用”这些问题进行了系统性的研究和解答。

（2）强度与位置结合分析。王玉琢等[13]基于上述计算方法，以及后来出现的自然语言处理领域的十大算法为基础，涉及全文引用次数以及引用的位置，将十大算法进行比较，得出每一套算法相较于其他算法的不同之处。“全文本文献计量分析”沙龙上提出[14]，需将引文位置分析和引文强度分析相结合，并对其全文中的分布进行可视化直观展示，寄于此希望对引用行为进行更直观的解读。

3.3 引用语境分析

主要涉及内容词和线索词两个概念，内容词主要包含信息的核心文本内容，而线索词只需要表示作者引用时的情感以及状态。内容词与线索词二者共同组成了引用语境分析的基础。对于引用语境分析，主要涉及词语选用、情感分析以及引用动机行为分析 3 个方面。

3.3.1 词语选用

在引用过程中，不同学科作者在引用时所选用的词语不同。这就引出了用词偏好与引用关系之间的探讨，就这一问题，HYLAND[15]表明，文章所用的动词会根据学科喝引用语境的不同而发生明显的动词使用差异，并且人文社科的学者比其他学科的使用的引文更多，且更倾向报告动词。KIM 等[16]利用引文内容和选用词语相似程度来表明作者主题词中的可能的隐性关系。胡志刚等[17]对引用语境分析中的连接词语，代词以及行为动词进行了分析通过对引用过程中的线索词进行收集，统计以及分析，对于使用词语的语境，辨明不同词语引用的功效分类，以及对引用语境进行模型构建起到了促进作用。ROSATI[18]基于语言学分析语料库，以揭示属于每种引文类型的独特且具有统计学意义的语言结构。

3.3.2 情感分析

引用与情感密切相关，在引文上下文情感分析方面，可以分为分类研究，方法研究以及模型构建。

（1）情感分析分类研究。分类 MUPPIDI 等[19]强调了对论文中引用情绪的正极性和负极性进行分类等问题。分类 AMJAD 等[20]通过使用 VerbNet 本体、自然语言处理和 4 种不同的机器学习算法对引文文本中的动词进行语义分析，并为研究人员提出了一种面向动词的引文情感分类。分类 VYAS 等[21]对于引文分析开发了 8 个包含引文句子的数据集，数据集由手动注释为 3

DOI：10.13998/j.cnki.issn1002-1248.23-0428

个情绪极性，即积极、消极和中性。

（2）情感分析方法研究。SMALL[22]在文章中提出，对引用语境的分析，需要用一种广义意义上的情感概念来进行分析，并要尊重涉及引文的不同态度和情感倾向。该研究采用语料库以及语言学等研究方法，对引文语境中出现的线索词集情况进行对比分析。IKRAM 等[23]提出了一种识别引文情绪的新方法，以揭示学术大数据中的隐藏模式。ALJUAID 等[24]提出了一种基于文本引文情感分析的二元分类方法，该方法有效促进了相关技术的发展。

（3）情感分析模型研究。UMER 等[25]在公开可用的引文情感数据集上使用机器学习模型结合基于频率和基于预测的特征表示技术的模式。MERCIER 等[26]根据研究结果提出了一个端对端的可训练多任务模型，该模型涵盖了利用域外数据集克服数据稀缺性的情绪和意图分析。

3.3.3 引用动机及引用行为分析

研究引用行为，可以利用实证调查和相关数据分析去验证之前引文分析得出的结论。一般将引用动机与引用行为的研究相结合，将之前通过不同研究方法得出的引用动机进行分析、归纳和合并，得出若干个重要的和较为完整的引用动机再去验证其有效性。首先要明确所依据的动机和行为分类标准，其次要确定研究方法，再确定研究对象以进行引用动机和行为分析。

（1）分类研究。对于这一分类问题，有许多学者做出了自己的判断。TEUFEL 等[27]的研究中，将学术引文进行分类，以此来对科学论证呼吁结构与引用的相互关系进行分析。并在此基础上体力了引用动机分类体系。SHEN 等[28]将引用分为“真引用”和“假引用”，倡导在进行文献计量分析时，更应该关注文章中出现的“真引用”。STEVENS 等[29]将引用文献的动机分为 15 种，其中包括对前人的研究表达尊敬、尊重同行的研究成果、研究方法以及研究工具的选取与比较、为全文提供知识背景、对以前研究的修正和个人意见的表达、对前期研究结果的批评等共 15 种引用动机分类。胡志刚[14]将引用动机和行为的数据分成两种：一是基于引用内容的数据来源；二是基于全文引文分析的数据来源。

（2）方法研究。对于引用动机的研究，使用的研究方法一般包括问卷调查法以及内容分析或者文献计量的方法来研究引用动机。①问卷调查法。WARRAICH 等[30]为了研究 LIS 专业毕业生的引用动机及其行为，进行了一项线上问卷调查。邱均平等[31]的研究说明，引用动机和行为并不以一成不变的，利用调查问卷进行数据手机，科研人员之间的相互影响对引用动机和行为产生一定的影响。②计量分析法。LIU 等[32]应用文本相似度算法分析科学家的三角引文行为。刘运梅等[33]同样利用三角引文原理，基于内容标注，通过内容标注的方式深入挖掘其中隐藏的引用动机。

（3）对象研究。引用动机以及引用行为的研究并不仅仅局限于期刊论文的领域，也可扩展到其他学术成果形式。LI 等[34]对于专利引用动机及行为做出研究，并提出 4 种假设并进行验证，以探究专利引用过程中的动机以及行为问题。

3.4 全文本引文分析的主要应用

全文本引文分析在文献计量学、科学学、科学评价等领域具有广泛的应用。

3.4.1 全文本引文分析应用于跨学科研究

（1）交叉测度研究。章成志等[35]对全文数据库中的不同学科的交叉情况进行引文内容分析测度。王芳等[36]对超过 50 本情报学领域杂志中的近 1 000 条理论进行频次统计分析，并分析了各个理论的来源学科，对中国情报学的理论来源分析起到了重要作用。张慧玲等[37]使用引文来对交叉学科杂志进行识别并对其识别方法进行研究，从跨学科领域特性的研究维度出发，对期刊之间的学科进行了共现分析与研究。徐庶睿等[38]对学科交叉成都进行了点面程度研究，探讨广义的学科主题交叉与侠义的学科主题交叉是由相关联。此后[39]在前面研究成果基础上，对引文内容和词语进行抽取收集，以探讨交叉度计算方法，并根据计算方法设置学科交叉度测度指标，以进行量化分析。梁镇涛等[40]基于引文角度出发，研究某一领域的学科交叉

的发展路径，并以眼球追踪领域为例进行了实证研究。张琳等[8]为了探究学科交叉测度的多样性指标，基于论文的不同章节的引文标注位置对学科交叉度进行分析，结果表明引言、讨论、方法和结果 4 个章节的学科交叉程度依次降低，并在此基础上提出了加权 TD 学科交叉测度指标，对学科交叉度测度综合指标进行了更新。HUANG 等[41]基于引文分析和语义分析提出了一个新的衡量跨学科互动的框架。

（2）知识流动研究。除了对交叉情况和交叉度进行研究，还可以通过全文本引文分析方法研究知识流动情况。刘丽帆等[42]采用引文内容分析法，以 PLoS 平台的开放学术论文全文数据为例，选取其中的引文句为研究对象。自动抽取引文句中包含的知识（即术语），再结合学科以及位置信息，揭示跨学科知识流动现象。孔玲等[43]提出了针对国内图书情报专业的学科交叉全文本引文的内容标注和联系分析框架，立足于分析引用强度、情感、位置及对象关联分析情报学跨学科被引行为，并在此基础上比较情报学跨学科知识扩散学科差异。张艺蔓等[44]结合引文内容分析和全文本引文分析，设置了施引、原始以及参考文献集这 3 个文献集，以探索学科内部与学科间知识流动的趋势[45]。商宪丽等[46]利用自然语言处理技术，从交叉领域文献题录数据、参考文献文本信息以及引文上下文中抽取知识短语，深入揭示交叉领域学科中“知识吸纳－知识内化”的知识整合过程。

（3）对外影响研究。CHEN 等[47]衡量了中国图书情报学对外部学科的学科影响程度。LEE 等[48]基于韩国角度出发，探究人力资源开发对其他学科的影响以及其他学科对人力资源开发研究的影响，来探索人力资源开发的跨学科性质，对其他学科的交叉学科研究有借鉴作用。

3.4.2　全文本引文分析应用于科学评价

目前对学术研究人员进行评价一般使用所发文章的 IF 值作为判断依据，但仍有各种研究试图找寻更合理合适的科学评价方式。

（1）学术评价。GILIAREVSKII 等[49]梳理了近代中国引入的科学活动成果评价的相关变化，以及世界主要国家对各国出版物在索引中的不平衡进行讨论，指出需要建立自己国家的评价指标以排除国外学术的优先权。在此基础上，迫切需要建立符合国家国情的，有自主权的，全面可靠的论文学术评价体系。CHEN 等[50]使用 Tobit 回归模型探讨跨学科成分的引用强度与论文科学影响力之间的关系。DAUD 等[51]根据引文的上下文对文章进行科学评估，得到基于上下文的文章影响因子（CBAIF），该因子不仅考虑引用的积极、消极或中立背景，还涉及引用和被引用作者的利益冲突关系。杨思洛等[52]将全文本分析方法融入传统二维评价模型，融合出一个新的指标，命名为学术和社会影响力指标，并在此基础上提出了新的论文影响力评价模型[53]。

（2）专利评价。彭爱东[54]、闫雪等[55]均对专利引文分析在企业竞争情报中的具体应用做出了探究，并进行了以联想企业为例的实证研究。王文平等[56]则对美国军工巨头在美国专利商标局中授权的专利引用情况统计分析，并探究专利引文与军工领域之间的关系。赵黎明等[57]对专利引文分析在知识转移机制研究中的应用做出探究，并对二者之间关系进行了深入探讨[58]，除此之外，LI 等[34]分析了审查人员和发明人 / 申请人的不同引用动机，并确定了 4 个不同假设以确定专利中的科学联系。

（3）作者评价。AN 等[59]从位置和内容两个角度来进行分析，探讨高被引作者的特点，并对不同位置进行具体分析，构建了不同位置的作者引用网络。GASPARYAN 等[60]综合了一些广泛使用的作者影响指标，并得到它们的最佳使用场景。ZHENG 等[61]基于 z 指数建立了一个基于主题作者引文的三维人才评价模型，并提出了 ZAS 指数来评价学科内部不同研究主题的学者。宋艳辉[62]基于引文分析与人力资本的计算，探究不同方法在人才评价体系中的应用，以构建更好指标更全面直观的评价人才。

3.4.3　全文本引文分析应用于在其他领域

关于未来的研究方向，ANDERSON[63]针对某一篇组织学习领域的文章进行引用分析，对其中的正面引用和负面引用均做了相关分析，并在此基础上确定了

DOI：10.13998/j.cnki.issn1002-1248.23-0428

未来研究的几个重要方向。关于如何利用全文本引文分析阅读领域文献，GIPP 等[64]提出引文邻近分析法（CPA）方法来对相近文章进行识别，可以在此基础上快速找到研究相关文献。关于如何更好判断作者共被引，JEONG 等[65]在传统方式作者共引分析（ACA）上，考虑作者的引用内容来衡量共同引用作者之间的相似性。关于引用价值，彭秋茹等[66]在引用强度、位置、情感等角度基础上，研究出信息计算引文指标以深入探讨引用的价值，得出专家普遍更认可结果位置的引用。张汝昊等[67]将引用语义和语境特征相互融合，基于施引动机的相似性这一本质层面优化引文耦合强度，提出一种融合引用语义和语境特征的作者引文耦合分析法。除了学术方面，全文本引文分析还可以泛化到其他领域，例如基于大数据挖掘技术，探究全文本视野下上海市政务新媒体的内容生产和传播特征[68]。或以微博社交媒体平台上的发布内容为数据来源，讨论微博写作中的新闻概念问题[69]。禹卫华[70]还在此基础上延探究全文本方法与传统的新闻研究方法的差异。

4 研究述评

4.1 全文本引文分析研究内容

（1）引用位置。总体而言，目前的研究结果显示，无论是定性还是定量，引文在全文中所处位置和引文的功能是相关的，一般来说处于文章重要结构的引文，在后续的评价体系以及涉及的学科交叉测度中所起到的作用就越大[9]。全文本引文分析除了常用于学术性文章的分析，也可以往外扩展到学术性著作，未来也可以再继续扩大适用范围例如专利、课题申报等。

（2）引用强度。引用强度研究往往不是单独出现的，一般都与引用位置研究密切相关。即使同一篇文章引用次数相同，但引用位置不同，引用强度也会有所不同。除此之外，还有关于引用次数与强度的计算方法，目前这方面的研究较少，有待进一步扩展。

（3）引用语境。在引用词语研究中，研究者更倾向于使用线索词，这可能与线索词的定义内涵以及属性有关，和内容词不同，线索词只表示情感和状态，与写作时的研究对象和研究内容没有太大关联，从这点来看，线索词可以作为引用语境分类的相关判断标准。对于内容词的进一步研究还需要后来者深入；目前对于引用情感分析，主要集中在情感分类以及情感分类模型等研究方面，今后随着计算机以及人工智能技术的进一步发展，机器学习深度学习等计算机研究技术及方法会与引用情感分析研究进一步融合，推动该领域的进一步发展；关于引用行为需要注意的是，作者的引用行为并不一定完全代表其引用动机的初衷，引用行为是一个可以清晰查看和量化的研究对象，但研究动机更多地偏向于一种思想过程，如何增加引用动机偏向的可靠性以及探明学术引用行为发生作用机制原理，还需要进一步的研究和论证。

4.2 全文本引文分析主要应用

全文本引用分析在跨学科研究方面，主要集中在测度方法、模型与指标的精进上，对于交叉学科研究来说，不仅要利用引文分析，还要结合其他计量学指标和测度方法，以便更好的对交叉学科进行测度；针对学术科学评价的讨论一直在进行中，中国以中科院为首一直在尝试降低国外期刊对中国学术界的门槛设置以及软影响。如何设置出合理的，符合中国国情的学术科学评价体系及指标，还需要进一步的深入研究，除此之外，还要注意社交媒体快速发展情况下网络媒体的引用问题；除了跨学科和科学评价上的应用以外，全文本引文分析还在其他领域有所发展，对某一学科发展趋势判断，深入挖掘线上数据等方面有所建树。

4.3 中外全文本分析的差异比较

（1）引用位置。国内前期侧重于全文本引用位置定性分析，主要用于判断不同学术发表形式中的引用位置是否与引文功效有相关关系；国外侧重于引用位置定量分析，多利用数学方法和公式对引用位置的具体功效进行定量化计算。近期国内也逐渐转变为定量分析为主，相继产出较为系统的可靠的定量量度方法。

（2）引用强度。引用强度分析主要分为单纯引用

强度以及结合引用位置的分析。通过对文献作者关系进行梳理，发现 DING 教授团队与国内章成志教授团队有合作关系并对该方向研究较深，不论单强度还是与位置结合分析均产出了一系列相关文献，促进了引用强度这一概念的推广。

（3）引用语境。词语选用方面，国外学者对词语的用词偏好和引用之间的关系进行较多研究，国内该方向以胡志刚教授团队研究较深入。且均侧重于线索词的分析研究，内容词由于其自身特性研究较少；情感分析方面，国外研究较为全面丰富，国内研究较少，目前侧重于情感分析模型的设置和训练上，依托大数据分析、机器学习、深度学习等计算机技术；引用动机和行为方面，国内外对其研究方向大致一致，对于动机和行为分类以及方法目前较为统一，且目前有向除文章以外的其他学术发表模式推广的趋势。

（4）引文应用。在学科交叉测度中，国外研究人员开始较早基础扎实，但近几年中国以张琳教授等人为代表，在前人基础上进行改革创新，推出较为完备的学科交叉测度综合指标体系，这可能与中国推行交叉学科体系制度建设有相关关系。在科学评价上，国外侧重加入全文本引文分析的评价模型建设，以及作者评价指标体系较为丰富，国内也提出了相关创新评价指标，但相比较于国外，对于专利评价研究较多。

总的来说，国外对于全文本引文分析的研究开始较早，但国内研究在前人的基础上也有较为丰厚的产出。目前国内外均重视与全文本引文分析与先进计算机技术结合，力求促进该领域发展。

5 全文本引文分析研究展望

5.1 全文本引文分析研究考虑的因素

全文本引文分析目前的研究仍然面临着许多问题和挑战，在展望全文本引文分析未来发展时，需要考虑以下 4 个方面。

（1）数据量和数据质量。全文本引文分析，需要借助全文本开放数据库的大量数据来进行数据支持，对于不同的开放数据网站，数据的质量有所不同，数据量的大小也有所不同。此外，目前的全文开放性数据库集中于医学生物等领域，对于社会学科的投入较少。目前各大出版集团均对旗下各杂志陆续开通 OA 网络首发等模式，数据量的持续增长是可以预见的。后续是否可以跨数据库进行研究，还需要进一步的分析。

（2）方法精度和效率。对于同一份数据，采用不同的计算方法得到的结果也有所不同。目前对于全文本引文分析的处理方法，主要借助于计算机领域的方法借用，目前还没有比较普遍可靠的属于本专业的成熟的计算机分析模型，需要进一步研究以适应本学科的发展，以提高全文本引文分析的精度和效率。

（3）跨语言引用分析。虽然当前的期刊文章语言一般以英语为主，但仍然涉及其他语言的情况，在分析引用情感动机等方面，需要根据目标文章所使用的语言来进行适当的调整。跨语言引用分析往往会面临许多难点和挑战，如语言差异、词汇差异等。英文需要考虑词语时态词性等问题，对于中文词语而言则需要考虑词语的断句以及多义词等复杂问题，还需将各不同语言专有名词一一对应。这些还只是跨语言分析需解决问题的冰山一角。需要后续探索更加准确和有效的跨语言引用分析方法和算法。

（4）可视化和交互性提升。Cite Space、Vos Viewer 等软件的产生和推广，导致可视化需要发生增长，VR 等技术的爆发导致交互性需要也提上日程。全文本引文分析需要与这一技术趋势相结合，在以往可视化的基础上，加入全文本引文分析的指标和可视化图形等，以求能够更直观的体现全文本引用分析结果。研究更加优秀的可视化和交互性处理方法，开发软件，以提高全文本引文分析的可用性和易用性。

5.2 全文本引文分析研究的未来趋势

基于以上分析，对于全文本引文分析未来的发展趋势以及可能的发展方向，做出来如下的可能预测。

（1）引用语义分析。当前的全文本引文分析主要集中在引用位置分析和引用强度分析上，而引用语义分析还相对较少。引用语义分析是指通过分析引用文

DOI：10.13998/j.cnki.issn1002-1248.23-0428

献中的语义信息，了解引文与被引文之间的相关性。随着自然语言处理技术的不断发展以及相关计算机编译方法的更新，引用语义分析或将成为全文本引用分析的一个重要方向。

（2）应用领域扩展。当前的全文本引用分析主要应用于学术研究领域，如学科发展趋势、学术合作关系、学术评价等方面。在今后的研究中，全文本引用分析可以扩展到更广泛的领域，如商业、经济学、医学技术等重要领域。

（3）大数据挑战。随着互联网技术和信息技术的迅猛发展，数据量呈爆炸式增长。全文本引用分析在面对海量数据时，需要处理大量的数据，挑战也随之而来。未来该研究领域的学术人员，是否能快速跟进技术的发展以促进该研究领域的拓展，又是否引进计算机相关人才，或改进本科及研究生的修习课程安排改革，需要今后进一步的研究和讨论。

（4）数据隐私保护。在进行全文本引文分析时，涉及大量的个人信息。如何保护个人数据隐私，是全文本引文分析需要解决的一个重要问题。

（5）实时性分析。当前的全文本引文分析主要基于静态数据，如论文、专利等，而实时性数据的引用分析还处于初级阶段。未来，全文本引文分析需要更好地适应实时数据的分析需求。

6 结　语

综上所述，目前全文本引文的分析主要集中于引用位置分析、引用强度分析以及引用语境分析上，细节上来说，情感分析热度较高。在应用方面，跨学科分析利用以及用于科学评价方面较为常见。面对诸多挑战，全文本引文分析仍然具有较大的发展潜力。通过对全文本引文分析现状，应用以及挑战进行梳理，希冀对后续研究提供参考和借鉴。

参考文献：

[1] 周慧娴. 全文本分析视角下的科学研究范式研究——以图书情报学为例[D]. 南京: 南京理工大学, 2020.

ZHOU H X. Research on scientific research paradigm from the perspective of full text analysis[D]. Nanjing: Nanjing University of Science and Technology, 2020.

[2] 何荣利, 魏洪善. 引文在论文中的分布和被引用内容的调查与分析[J]. 图书情报工作, 2000, 44(2): 26-29.

HE R L, WEI H S. The distribution of citations in papers and the investigation and analysis of the substance quoted[J]. Library and information service, 2000, 44(2): 26-29.

[3] 李燕. 关于引文分布和引文效用的调查分析[J]. 现代情报, 2010, 30(6): 145-147.

LI Y. Investigation and analysis on the distribution of citations and effectiveness of citations[J]. Journal of modern information, 2010, 30(6): 145-147.

[4] 刘茜, 王健, 王剑, 等. 引文位置时序变化研究及其认知解释[J]. 情报杂志, 2013, 32(5): 166-169, 184.

LIU Q, WANG J, WANG J, et al. A study on timing variations in citations position and its cognitive interpretation[J]. Journal of intelligence, 2013, 32(5): 166-169, 184.

[5] HALEVI G, MOED H F. The thematic and conceptual flow of disciplinary research: A citation context analysis of thejournal of informetrics, 2007[J]. Journal of the American society for information science and technology, 2013, 64(9): 1903-1913.

[6] 胡志刚. 全文引文分析: 理论、方法与应用[M]. 北京: 科学出版社, 2016.

HU Z G. Full-text citation analysis: Theory, method and application[M]. Beijing: Science Press, 2016.

[7] 章成志, 王玉琢, 卢超. 学术专著引用行为研究——基于引文内容特征分析的视角[J]. 情报学报, 2017, 36(3): 319-330.

ZHANG C Z, WANG Y Z, LU C. Citing behavior of academic monographs - Perspective based on character analysis of citation content[J]. Journal of the China society for scientific and technical information, 2017, 36(3): 319-330.

[8] 张琳, 刘冬东, 吕琦, 等. 论文学科交叉测度研究：从全部引文到章节引文[J]. 情报学报, 2020, 39(5): 492-499.

ZHANG L, LIU D D, LYU Q, et al. Interdisciplinarity measurement in publications: From full reference analysis to sectional reference analysis[J]. Journal of the China society for scientific and technical infor-

mation, 2020, 39(5): 492-499.

[9] 卢超, 章成志, 王玉琢, 等. 语义特征分析的深化——学术文献的全文计量分析研究综述[J]. 中国图书馆学报, 2021, 47(2): 110-131.
LU C, ZHANG C Z, WANG Y Z, et al. Strengthened analyses of semantic features: Review of full-text bibliometrics of academic documents[J]. Journal of library science in China, 2021, 47(2): 110-131.

[10] DING Y, LIU X Z, GUO C, et al. The distribution of references across texts: Some implications for citation analysis[J]. Journal of informetrics, 2013, 7(3): 583-592.

[11] 胡志刚, 陈超美, 刘则渊, 等. 从基于引文到基于引用——一种统计引文总被引次数的新方法[J]. 图书情报工作, 2013, 57(21): 5-10.
HU Z G, CHEN C M, LIU Z Y, et al. From counting references to counting citations: A new way to calculate the total cited times of references[J]. Library and information service, 2013, 57(21): 5-10.

[12] HU Z G, LIN G G, SUN T A, et al. Understanding multiply mentioned references[J]. Journal of informetrics, 2017, 11(4): 948-958.

[13] 王玉琢, 章成志. 考虑全文本内容的算法学术影响力分析研究[J]. 图书情报工作, 2017, 61(23): 6-14.
WANG Y Z, ZHANG C Z. Using full-text to analyse academic impact of algorithms[J]. Library and information service, 2017, 61(23): 6-14.

[14] 王佳敏, 李信, 刘齐进. 全文本文献计量分析学术沙龙综述[J]. 信息资源管理学报, 2018, 8(4): 119-125.
WANG J M, LI X, LIU Q J. A review of the academic salon on full-text bibliometric analysis[J]. Journal of information resources management, 2018, 8(4): 119-125.

[15] HYLAND K. Academic attribution: Citation and the construction of disciplinary knowledge[J]. Applied linguistics, 1999, 20(3): 341-367.

[16] KIM H J, JEONG Y K, SONG M. Content- and proximity-based author co-citation analysis using citation sentences[J]. Journal of informetrics, 2016, 10(4): 954-966.

[17] 胡志刚, 孙太安, 王贤文. 引用语境中的线索词分析——以 Journal of Informetrics 为例[J]. 图书情报工作, 2017, 61(23): 25-33.
HU Z G, SUN T A, WANG X W. Citation context analysis based on cue words: A case study of journal of informetrics[J]. Library and information service, 2017, 61(23): 25-33.

[18] ROSATI D. Citations are not opinions: A corpus linguistics approach to understanding how citations are made[J]. arXiv: 2104.08087,2021.

[19] MUPPIDI S, GORRIPATI S K, KISHORE B. An approach for bibliographic citation sentiment analysis using deep learning[J]. International journal of knowledge-based and intelligent engineering systems, 2021, 24(4): 353-362.

[20] AMJAD Z, IHSAN I. VerbNet based citation sentiment class assignment using machine learning[J]. International journal of advanced computer science and applications, 2020, 11(9): 621-627.

[21] VYAS V, RAVI K, RAVI V, et al. Article citation study: Context enhanced citation sentiment detection[J]. 2020: arXiv: 2005.04534, 2020.

[22] SMALL H. Interpreting maps of science using citation context sentiments: A preliminary investigation[J]. Scientometrics, 2011, 87(2): 373-388.

[23] IKRAM M T, AFZAL M T. Aspect based citation sentiment analysis using linguistic patterns for better comprehension of scientific knowledge[J]. Scientometrics, 2019, 119(1): 73-95.

[24] ALJUAID H, IFTIKHAR R, AHMAD S, et al. Important citation identification using sentiment analysis of in-text citations[J]. Telematics and informatics, 2021, 56: 101492.

[25] UMER M, SADIQ S, MISSEN M M S, et al. Scientific papers citation analysis using textual features and SMOTE resampling techniques[J]. Pattern recognition letters, 2021, 150: 250-257.

[26] MERCIER D, RIZVI S T R, RAJASHEKAR V, et al. Utilizing out-domain datasets to enhance multi-task citation analysis[J]. arXiv: 2202.10884, 2022.

[27] TEUFEL S, SIDDHARTHAN A, TIDHAR D. An annotation scheme for citation function[C]// Proceedings of the 7th SIGdial Workshop on Discourse and Dialogue - Sig-DIAL' 06. Morristown, NJ, USA: Association for Computational Linguistics, 2006: 80-87.

[28] SHEN Z S, YANG L Y, PEI J S, et al. Interrelations among scientific fields and their relative influences revealed by an input-output analysis[J]. Journal of informetrics, 2016, 10(1): 82-97.

[29] STEVENS M, GIULIANO V, GARFIELD E. Can citation indexing be automated? [C]. Washington: Symposyum on Statistical Assoc Methods for Mechanized Documentation, 1964.

DOI：10.13998/j.cnki.issn1002-1248.23-0428

[30] WARRAICH N F, ALI I, TAHIRA M, et al. Factors influencing the citation behavior of Pakistani novice LIS researchers[J]. Journal of librarianship and information science, 2023, 55(2): 434-443.

[31] 邱均平, 陈晓宇, 何文静. 科研人员论文引用动机及相互影响关系研究[J]. 图书情报工作, 2015, 59(9): 36-44.

QIU J P, CHEN X Y, HE W J. Study on paper citation motivations and mutual influence of researchers[J]. Library and information service, 2015, 59(9): 36-44.

[32] LIU Y M, CHEN M. Applying text similarity algorithm to analyze the triangular citation behavior of scientists[J]. Applied soft computing, 2021, 107: 107362.

[33] 刘运梅, 张帅, 司湘云, 等. 基于内容标注的三角引用动机研究方法探析[J]. 图书情报工作, 2021, 65(10): 48-55.

LIU Y M, ZHANG S, SI X Y, et al. Investigation and research on triangular citation moti-vation based on content annotation[J]. Library and information service, 2021, 65(10): 48-55.

[34] LI R, CHAMBERS T, DING Y, et al. Patent citation analysis: Calculating science linkage based on citing motivation[J]. Journal of the association for information science and technology, 2014, 65 (5): 1007-1017.

[35] 章成志, 徐庶睿, 卢超. 利用引文内容监测多学科交叉现象的方法与实证[J]. 图书情报工作, 2016, 60(19): 108-115.

ZHANG C Z, XU S R, LU C. Using citation content for interdisciplinary phenomenon[J]. Library and information service, 2016, 60 (19): 108-115.

[36] 王芳, 陈锋, 祝娜, 等. 我国情报学理论的来源、应用及学科专属度研究[J]. 情报学报, 2016, 35(11): 1148-1164.

WANG F, CHEN F, ZHU N, et al. Theories of information science in China: Sources, uses and discipline exclusive degrees[J]. Journal of the China society for scientific and technical information, 2016, 35(11): 1148-1164.

[37] 张慧玲, 许海云, 刘春江. 学科交叉期刊的识别方法研究[J]. 情报学报, 2018, 37(11): 1140-1153.

ZHANG H L, XU H Y, LIU C J. Study on recognition methods for interdisciplinary periodicals [J]. Journal of the China society for scientific and technical information, 2018, 37(11): 1140-1153.

[38] 徐庶睿, 卢超, 章成志. 学科交叉度的点面关系研究[J]. 图书馆论坛, 2017, 37(10): 64-70.

XU S R, LU C, ZHANG C Z. A study of interdisciplinary degree: From macro-to micro-scopic[J]. Library tribune, 2017, 37(10): 64-70.

[39] 徐庶睿, 章成志, 卢超. 利用引文内容进行主题级学科交叉类型分析[J]. 图书情报工作, 2017, 61(23): 15-24.

XU S R, ZHANG C Z, LU C. Using citation contents for the interdisciplinary type analysis at a topical level[J]. Library and information service, 2017, 61(23): 15-24.

[40] 梁镇涛, 巴志超, 徐健. 基于引文的跨学科领域发展路径分析——以眼动追踪领域为例[J]. 图书情报工作, 2019, 63(23): 65-78.

LIANG Z T, BA Z C, XU J. Development path of interdisciplinary field based on citation analysis: A case study of eye tracking[J]. Library and information service, 2019, 63(23): 65-78.

[41] HUANG L, CAI Y J, ZHAO E D, et al. Measuring the interdisciplinarity of information and library science interactions using citation analysis and semantic analysis[J]. Scientometrics, 2022, 127(11): 6733-6761.

[42] 刘丽帆, 张恒, 章成志. 基于学术文献引文内容的跨学科知识流动研究[J]. 情报理论与实践, 2022, 45(6): 24-31, 47.

LIU L F, ZHANG H, ZHANG C Z. Interdisciplinary knowledge flows based on citation contents of academic papers[J]. Information studies: Theory & application, 2022, 45(6): 24-31, 47.

[43] 孔玲, 张卫, 王东波, 等. 全文本引文视角下我国情报学跨学科知识扩散研究[J]. 文献与数据学报, 2022, 4(4): 110-128.

KONG L, ZHANG W, WANG D B, et al. Interdisciplinary knowledge diffusion in information science of China from the perspective of full-text citation[J]. Journal of library and data, 2022, 4(4): 110-128.

[44] 张艺蔓, 马秀峰, 程结晶. 融合引文内容和全文本引文分析的知识流动研究[J]. 情报杂志, 2015, 34(11): 50-54, 49.

ZHANG Y M, MA X F, CHENG J J. Research of knowledge flows based on citation content analysis[J]. Journal of intelligence, 2015, 34(11): 50-54, 49.

[45] 尹莉. “极性”概念在引文分析中应用的一个实证研究[J]. 情报杂志, 2017, 36(8): 124-130, 143.

YIN L. Empirical research on applying concept of polarity to citation analysis[J]. Journal of intelligence, 2017, 36(8): 124-130, 143.

[46] 商宪丽, 王施运, 操玉杰, 等. 基于引文内容分析的交叉领域知识内化过程研究[J]. 情报科学, 2023, 41(2): 118-125.
SHANG X L, WANG S Y, CAO Y J, et al. Analyzing the knowledge internalization process of interdisciplinary field based on citation content[J]. Information science, 2023, 41(2): 118-125.

[47] CHEN C F, LI Q A, CHIU K, et al. The impact of Chinese library and information science on outside disciplines: A citation analysis [J]. Journal of librarianship and information science, 2020, 52(2): 493-508.

[48] LEE J Y, BAEK P. Exploring the interdisciplinary nature of HRD using citation analysis[J]. The Korean journal of human resource development quarterly, 2022, 24(3): 1-35.

[49] GILIAREVSKII R S, MELNIKOVA E V. Rejection of the priority of international science citation indices in the evaluation of scientific activity results in China[J]. Scientific and technical information processing, 2020, 47(3): 194-199.

[50] CHEN S J, SONG Y H, QIU J P, et al. The effect of interdisciplinary components' citation intensity on scientific impact[J]. Libr hi tech, 2021, 39: 1084-1096.

[51] DAUD A, GHAFFAR S, AMJAD T. Citation count is not enough: Citation's context -based scientific impact evaluation [J]. IEEE transactions on computational social systems, 2022(99): 1-7.

[52] 杨思洛, 聂颖. 结合全文本分析的论文影响力评价模型研究[J]. 现代情报, 2022, 42(3): 133-146.
YANG S L, NIE Y. Research on evaluation model of papers' influence combined with full-text analysis[J]. Journal of modern information, 2022, 42(3): 133-146.

[53] 伍军红, 肖宏, 孙隽, 等. 从单篇论文引证视角改进学术期刊评价方法及其实证研究[J]. 中国科技期刊研究, 2022, 33(7): 957-964.
WU J H, XIAO H, SUN J, et al. Improving the evaluation method of academic journals from citation perspective of single paper and its empirical research[J]. Chinese journal of scientific and technical periodicals, 2022, 33(7): 957-964.

[54] 彭爱东. 专利引文分析在企业竞争情报中的应用[J]. 情报理论与实践, 2004, 27(3): 276-278.
PENG A D. The application of patent citation analysis in enterprise competitive intelligence[J]. Information studies: Theory & application, 2004, 27(3): 276-278.

[55] 闫雪, 彭爱东, 朱小聪, 等. 专利引文分析在企业竞争力研究中的应用——以联想集团为例[J]. 情报理论与实践, 2009, 32(1): 76-78, 89.
YAN X, PENG A D, ZHU X C, et al. Application of patent citation analysis in enterprise competency research[J]. Information studies: Theory & application, 2009, 32(1): 76-78, 89.

[56] 王文平, 刘云, 叶选挺. 专利引文分析在军工企业中的应用研究[J]. 兵工学报, 2009, 30(S1): 45-50.
WANG W P, LIU Y, YE X T. Research on the application of patent citation analysis in defense firms[J]. Acta armamentarii, 2009, 30(S1): 45-50.

[57] 赵黎明, 高杨, 韩宇. 专利引文分析在知识转移机制研究中的应用[J]. 科学学研究, 2002, 20(3): 297-300.
ZHAO L M, GAO Y, HAN Y. Application of patent citation analysis to the research of knowledge-transfer mechanism[J]. Studies in science of science, 2002, 20(3): 297-300.

[58] 闫昊. 中国专利引文分布特征分析与应用研究——以5G领域为例[D]. 太原: 山西大学, 2021.
YAN H. Analysis and application of Chinese patent citation distribution characteristics[D]. Taiyuan: Shanxi University, 2021.

[59] AN J, KIM N, KAN M Y, et al. Exploring characteristics of highly cited authors according to citation location and content[J]. Journal of the association for information science and technology, 2017, 68(8): 1975-1988.

[60] GASPARYAN A Y, YESSIRKEPOV M, DUISENOVA A, et al. Researcher and author impact metrics: Variety, value, and context[J]. Journal of Korean medical science, 2018, 33(18): 1-16.

[61] ZHENG Y Y, LIU S. Bibliometric analysis for talent identification by the subject-author-citation three-dimensional evaluation model in the discipline of physical education[J]. Libr hi tech, 2020, 40: 62-79.

[62] 宋艳辉. 知识计量学在人才评价中的应用——基于引文分析与人力资本测算的研究[J]. 重庆大学学报(社会科学版), 2013, 19(6): 89-93.
SONG Y H. Application research of knowmetrics for talents evaluation based on citation analysis and human capital measure[J]. Journal of Chongqing university (social science edition), 2013, 19(6):

DOI：10.13998/j.cnki.issn1002-1248.23-0428

89-93.

[63] ANDERSON M H, SUN P Y T. What have scholars retrieved from Walsh and Ungson (1991)? A citation context study[J]. Management learning, 2010, 41(2): 131-145.

[64] GIPP B, BEEL J. Citation proximity analysis (CPA) - A new approach for identifying related work based on co-citation analysis[C]. Rio de Janeiro: 12th International Conference on Scientometrics and Informetrics, 2009: 571-575.

[65] JEONG Y K, SONG M, DING Y. Content-based author co-citation analysis[J]. Journal of informetrics, 2014, 8(1): 197-211.

[66] 彭秋茹, 阎素兰, 黄水清. 基于全文本分析的引文指标研究——以 F1000 推荐论文为例[J]. 信息资源管理学报, 2019, 9(4): 82-88.

PENG Q R, YAN S L, HUANG S Q. Research on the citation indexes based on full-text analysis: Taking the F1000 articles as examples[J]. Journal of information resources management, 2019, 9(4): 82-88.

[67] 张汝昊, 袁军鹏. 融合引用语义和语境特征的作者引文耦合分析法[J]. 情报学报, 2022, 41(8): 796-811.

ZHANG R H, YUAN J P. Semantic-and contextual-based author bibliographic coupling analysis[J]. Journal of the China society for scientific and technical information, 2022, 41(8): 796-811.

[68] 王玲宁, 禹卫华. 全文本视野下政务新媒体的内容生产和传播特征——以"上海发布"为例[J]. 新闻界, 2017(9): 27-31.

WANG L N, YU W H. Characteristics of content production and dissemination of new government media from the perspective of full text - Taking "Shanghai Publishing" as an example[J]. Journalism and mass communication monthly, 2017(9): 27-31.

[69] 王甬平, 孙祥飞. 试论微博写作中的新闻概念——以《人民日报》法人微博的全文本分析为案例[J]. 中国广播电视学刊, 2014(2): 73-75.

WANG Y P, SUN X F. On the concept of news in Weibo's writing-A case study of Weibo, the legal person of People's Daily[J]. China radio & TV academic journal, 2014(2): 73-75.

[70] 禹卫华. 社交媒体全文本分析法刍议[J]. 新闻记者, 2015(12): 86-89.

YU W H. On full-text analysis of social media[J]. Journalism review, 2015(12): 86-89.

Review of Full Text Citation Analysis

LI Dian, WEN Tingxiao*, XU Linyong*

(School of Life Sciences, Central South University, Changsha 410006)

Abstract: [Purpose/Significance] Citation analysis is an important research field of biblio-metrics, which aims to investigate the citation behavior and academic influence of literature by analyzing the citation relationship of literature. Specifically, full-text citation analysis can help researchers understand the citation relationship and academic impact of the literature more accurately and comprehensively. This paper aims to clarify the development of full-text citation analysis, review the different progress of full-text citation research, and on this basis, attempt to analyze the challenges of full-text citation analysis and the possible development direction in the future in order to better promote the development of citation analysis of the full text. [Method/Process] Key words and search formulas were determined on the CNKI database and the Web of Science database, and then the retrieved results were screened and filtered according to the summary. This paper reviews the research results of full-text citation analysis at home and abroad, and summarizes the research contents of full-text citation analysis into citation location analysis, citation intensity analysis and citation context analysis. Citation context

analysis can be divided into word selection, emotion analysis, motivation and behavior analysis. At the same time, the application of citation analysis in interdisciplinary and scientific evaluation is analyzed. A summary of the differences in research content and application of full-text citation analysis at home and abroad was provided. [Results/Conclusions] The current research on citation analysis of the full text has an important connection with the chapter structure. The reference strength analysis is closely combined with the reference position analysis, and the focus is to develop new calculation methods. In the context of citation, the emphasis of citation analysis is emotional analysis. The application mainly focuses on the construction of interdisciplinary measurement and scientific evaluation system. This paper also predicts the challenges and future directions of citation analysis. The specific factors to be considered include data volume and data quality, method accuracy and efficiency, cross language reference analysis, and the improvement of visualization and interactivity. In spite of many challenges, the full text citation analysis still has great development potential. In the future, it is possible to make progress in citation semantic analysis, expansion of application field, big data analysis, data privacy protection and real-time analysis. This paper reviews the current situation, application and challenges of citation analysis, hoping to provide reference and insights for future research. At the same time, this study still has some shortcomings, based on the development of computer technology, it has some limitations in its discussion on the future development trend.

Keywords: full text analysis; citation analysis; emotional analysis; reference location analysis